

On Local RBF Approximation ^{*†}

T.Hangelbroek[‡]

September 29, 2009

Abstract

The purpose of this paper is to investigate RBF approximation with highly nonuniform centers. Recently, DeVore and Ron have developed a notion of the local density of a set of centers – a notion that permits precise pointwise error estimates for surface spline approximation. We give an equivalent, alternative characterization of local density, one that allows effective placement of centers at different resolutions.

1 Introduction

This brief article concerns local approximation results for radial basis function (RBF) approximation with the goal of effectively placing centers at varying resolutions. We consider RBF approximants of the form $x \mapsto \sum_{\xi \in \Xi} \phi(x - \xi)$, where the arrangement of centers, Ξ , may be highly nonuniform. A motivation for this set up is that centers may be placed strategically to treat defects in the target function. This becomes very important in high dimensions, where conventional quasi-uniform placement of centers is extremely costly; error estimates assuming a (small) fill distance h require a placement of $\mathcal{O}(h^{-d})$ centers; obtaining a comparable error with fewer centers is clearly desirable. To this end, we seek a method by which Ξ can be chosen to achieve a pointwise error that reflects the local arrangement of Ξ . In turn, this requires finding a useful measure of the *local density* of Ξ .

In [3], DeVore and Ron establish powerful local error estimates for kernel based approximation; along the way they develop a satisfactory notion of local density – the *majorant* – expressed as a function over a domain containing the centers (see (2) below). This function gives, roughly, the distance to the nearest unisolvent subset of Ξ . However, it also satisfies an extra condition of global compatibility: it may not grow or shrink too rapidly. This condition is not stated explicitly, rather

^{*}2000 *Mathematics Subject Classification*: 41A17, 26B35, 41A63, 42C15

[†]*Key words*: error estimates, nonlinear approximation, optimal approximation, radial basis functions, scattered data, thin-plate splines, surface splines, approximation order

[‡]Texas A&M University College Station, TX 77843, USA. Research supported by NSF Postdoctoral Research Fellowship.

it is contained in the definition, but it plays an essential role in their local error estimates. In this article we investigate this compatibility condition and give an equivalent condition that lends itself to effective placement of highly nonuniform centers.

Error estimates in [3] show that kernel approximation bears a strong similarity to univariate spline approximation, which exhibits local convergence in the sense that error decays rapidly over intervals where breakpoints are tightly spaced. Indeed, such results have long been known for spline quasi-interpolation, at least since [1]. If Q_T is the quasi-interpolation operator of order r associated with knots $T = (t_j)$, then [2, (4.18)] in conjunction with a theorem of Whitney [2, Theorem 4.2] tells us that for $x \in [t_j, t_{j+1}]$,

$$|f(x) - Q_T(f)(x)| \leq C_r \omega_{r+1}(f, t_{j+r} - t_{j-r+1}).$$

We note that the error from spline approximation at x depends only on the distribution of knots near x (the r^{th} *nearest neighbors*) and the smoothness of the target function in a neighborhood of x . This result can be attributed to the local nature of the B-splines basis (and its associated dual functionals).

Because kernels are often globally supported, and because a truly local basis similar to the B-splines seems to be out of the question, the distance to the nearest neighbors may be unsuitable for measuring local density (see Remark 3 below for evidence to the contrary). Kernel approximants exhibit *far field* effects, meaning that each kernel has a global influence, and the majorant of DeVore–Ron, via the global compatibility condition, penalizes remote, sparse density in an effort to mitigate such far field effects.

In the following section, we present an alternative, equivalent characterization of the majorant of DeVore and Ron, one that involves a global compatibility condition, and we give a self-contained development of their pointwise error estimates that explicitly uses this compatibility condition. In Section 3, we discuss how the compatibility condition may be used to place centers nonuniformly.

2 Local Estimates

In [3], DeVore and Ron construct a local approximation scheme using a simple measure of the local density of centers. This initial notion of density (given in Definition 1) is not suitable to capture far field effects, and the analysis of the scheme’s convergence eventually relies on a more refined notion of local density: the initial density’s *majorant*. In this section, we recast the DeVore–Ron result with a different but equivalent local density parameter, one that lends itself to efficient placement of centers. As in [3], we begin by giving the initial definition of local density, a function on \mathbb{R}^d which at each $\alpha \in \mathbb{R}^d$ indicates a radius sufficient to capture a K -stable local polynomial reproduction of order ℓ :

Definition 1 (Local Density). *Given a set of centers $\Xi \subset \mathbb{R}^d$, a local density (LD) $\rho : \mathbb{R}^d \rightarrow \mathbb{R}_+$ is a function with an associated local polynomial reproduction of precision ℓ . That is to say, there is a kernel $a : \Xi \times \mathbb{R}^d \rightarrow \mathbb{R} : (\xi, \alpha) \mapsto a(\xi, \alpha)$ so that the following hold:*

(Support) For $|\xi - \alpha| > \rho(\alpha)$, $a(\xi, \alpha) = 0$.

(Precision) For all $p \in \Pi_\ell$ we have $\sum_{\xi \in \Xi} a(\xi, \alpha) p(\xi) = p(\alpha)$.

(Stability) There is $K > 0$ such that $\sum_{\xi \in \Xi} |a(\xi, \alpha)| < K$ for all α .

This construction allows the surface spline, ϕ (also known as polyharmonic splines because they are the fundamental solution of the k -fold Laplacian Δ^k), to be approximated by a linear combination of nearby shifts. This is accomplished with a small, local error:

$$|\phi(x - \alpha) - \sum a(\xi, \alpha) \phi(x - \xi)| \leq C \rho(\alpha)^{2k-d} \left(1 + \frac{|x - \alpha|}{\rho(\alpha)}\right)^{-\nu}, \quad \nu := \ell + d - 2k,$$

and one can generate the approximant:

$$T_\Xi f(x) := \int_{\mathbb{R}^d} \Delta^k f(\alpha) \sum_{\xi \in \Xi} a(\xi, \alpha) \phi(x - \xi) d\alpha.$$

This leads to convenient pointwise error estimates:

$$|f(x) - T_\Xi f(x)| \leq C \int_{\mathbb{R}^d} |\Delta^k f(\alpha)| \rho(\alpha)^{2k-d} \left(1 + \frac{|x - \alpha|}{\rho(\alpha)}\right)^{-\nu} d\alpha. \quad (1)$$

Observe that the construction of the approximation operator T_Ξ depends only on a local polynomial reproduction $(\xi, \alpha) \mapsto a(\xi, \alpha)$. The estimate (1) holds for any LD with local polynomial reproduction a . In particular, it holds for any function ρ' with $\rho' \geq \rho$. With this in mind, it may seem tempting to use an optimally small ρ : the LD so that $\rho(\alpha)$ is the minimal radius around α that captures a K -stable polynomial reproduction of order ℓ . Sadly, because of the global nature of the kernels, this is unsuitable for producing precise estimates.

In [3] the LD is used to create a “majorant”:

$$H(x) = \sup_{y \in \mathbb{R}^d} \rho(y) \left(1 + \frac{|x - y|}{\rho(y)}\right)^{-r}. \quad (2)$$

In short, the coefficient kernel associated to the LD is used to create the approximant, but, in order to attack the estimate (1), an expression involving the majorant controls the error. The initial LD only plays an ancillary role: to construct the approximant via the coefficient kernel and to generate the majorant.

To be sure, the entire exercise could be repeated using only H and entirely without ρ . It is a simple task to show that H is itself an LD (since $H(x) \geq \rho(x)$ and the local polynomial reproduction $(\xi, x) \mapsto a(\xi, x)$ is a local polynomial reproduction for H as well). Moreover, the majorant of H is simply a constant multiple of H (a constant depending only on r). It follows that by replacing the initial LD ρ with its majorant H , one would obtain the same results.

Alternately, from the beginning one may impose the assumption that the LD is equivalent to its majorant (i.e., it is *self-majorizing*). This is a change of perspective: from the point of view that centers have been given outside of our control (and with the goal of remaining faithful to the local distribution of data by having error estimates reflecting the local density) to the setting where the spacing of centers is chosen to reflect regions of interest or to attack defects in the target function. We will choose centers having an LD that is correct from the start, having an extra condition designed to handle far field effects. The condition on the LD will be different from that of [3]; it is a slightly more easily verified property: *slow growth*. To proceed, we formalize both extra assumptions (slow growth and self-majorization) under the heading of global compatibility and take a moment to discuss their equivalence.

Definition 2 (Global Compatibility). *If there is a constant $C_{sg} > 0$ so that for every x and α , we have*

$$\rho(\alpha) \leq C_{sg}\rho(x) \left(1 + \frac{|x - \alpha|}{\rho(x)}\right)^{1-\epsilon}, \quad (3)$$

we say ρ exhibits $1 - \epsilon$ slow growth.

The function $\rho : \Omega \rightarrow \mathbb{R}_+$ exhibits self-majorization of order r if there is a constant $C_{sm} > 0$ so that for every x and y , we have

$$\rho(y) \geq C_{sm}\rho(x) \left(1 + \frac{|x - y|}{\rho(x)}\right)^{-r} \quad (4)$$

The equivalence of these two assumptions can be expressed formally:

Lemma 3. *If ρ satisfies the property of self-majorization (4) then it satisfies the property of slow growth (3) with $\epsilon = \frac{1}{r+1}$ and constant C_{sg} depending only on r and C_{sm} . Likewise, if ρ satisfies (3) then it satisfies (4) with $r = \frac{1-\epsilon}{\epsilon}$ and constant C_{sm} depending only on ϵ and C_{sg} .*

Proof. When $|x - \alpha| \geq \rho(x)$, (3) implies that $\rho(\alpha) \leq 2^{1-\epsilon}C_{sg}\rho(x)^\epsilon|x - \alpha|^{1-\epsilon}$, so

$$\frac{2^{\epsilon-1}}{C_{sg}}\rho(\alpha)^\epsilon \left(\frac{|x - \alpha|}{\rho(\alpha)}\right)^{\epsilon-1} \leq \rho(x)^\epsilon.$$

On the other hand, when $|x - \alpha| < \rho(x)$, $\rho(\alpha) \leq 2^{1-\epsilon}C_{sg}\rho(x)$, so ρ satisfies self-majorization with $C_{sm} = \min\left(\frac{2^{\epsilon-1}}{C_{sg}}, \left(\frac{2^{\epsilon-1}}{C_{sg}}\right)^{1/\epsilon}\right)$.

When $|x - \alpha| \geq \rho(x)$, (4) implies that $\rho(\alpha) \geq 2^{-r}C_{sm}\rho(x)^{1+r}|x - \alpha|^{-r}$, so

$$\frac{2^r}{C_{sm}}\rho(\alpha)^{1+r} \left(\frac{|x - \alpha|}{\rho(\alpha)}\right)^r \geq \rho(x)^{1+r}.$$

On the other hand, when $|x - \alpha| < \rho(x)$, $\rho(\alpha) \leq 2^{-r}C_{sm}\rho(x)$, so ρ satisfies slow growth with $C_{sg} = \max\left(\frac{2^r}{C_{sm}}, \left(\frac{2^r}{C_{sm}}\right)^{1/(1+r)}\right)$. □

Either of these extra assumptions on ρ are sufficient to obtain the error estimate in [3]:

Theorem 4 (DeVore Ron I). *Let $\ell > 2k - d + 1$. Suppose that ρ satisfies Slow Growth with $\epsilon > \frac{2k}{\ell}$. There is a constant C so that for $f \in C^{2k}(\mathbb{R}^d)$ having compact support,*

$$|f(x) - T_{\Xi}f(x)| \leq C\rho(x)^{2k}\|\Delta^k f\|_{\infty}$$

Proof. This follows by applying the growth assumption to (1) and writing $\gamma = 1 - \epsilon$ to obtain:

$$\begin{aligned} |f(x) - T_{\Xi}f(x)| &\leq C \int_{\mathbb{R}^d} |\Delta^k f(\alpha)| \rho(x)^{2k-d} \left(1 + \frac{|x - \alpha|}{\rho(x)}\right)^{\gamma(2k-d)} \left(1 + \frac{\frac{|x - \alpha|}{\rho(x)}}{\left(1 + \frac{|x - \alpha|}{\rho(x)}\right)^{\gamma}}\right)^{-\nu} d\alpha \\ &\leq C\rho(x)^{2k-d}\|\Delta^k f(\alpha)\|_{\infty} \int_{\mathbb{R}^d} \left(1 + \frac{|x - \alpha|}{\rho(x)}\right)^{2k-d-\ell+\ell\gamma} d\alpha \\ &\leq C\rho(x)^{2k}\|\Delta^k f(\alpha)\|_{\infty} \int_0^{\infty} (1 + R)^{2k-d-\ell+\ell\gamma} R^{d-1} dR. \end{aligned}$$

The second inequality follows by writing $(2k - d)\gamma - \nu(1 - \gamma) = (2k - d)\gamma - (\ell - 2k + d)(1 - \gamma)$. The convergence of the last integral is a consequence of the assumption $\gamma < 1 - \frac{2k}{\ell}$. \square

A further result from [3], is that functions of lower smoothness can also be treated with local error estimates. The operator T_{Ξ} is instrumental in obtaining low smoothness results, albeit indirectly. This is the point of [3, Theorem 5.3], which, for completeness, we rephrase in a simplified form as Theorem 5 in terms of the slow growth assumption. The lower order result is technically more complicated than that of Theorem 4. It is a common technique to use interpolation theory to obtain direct approximation results for functions of lower smoothness. DeVore and Ron use an argument of this type, that splits f into a rough but benign part, b , and a smooth part, g . This is done in a way that is not entirely straightforward, by controlling the size of b (and the smoothness of g) in a precise way to match the local density.

To discuss lower smoothness estimates, we first introduce fractional smoothness spaces. These can be expressed in numerous different ways: as Besov spaces, Triebel-Lizorkin spaces or (more familiarly) Hölder-Zygmund spaces. In the setting we consider, these are the same spaces. That is, we consider $F_{\infty,\infty}^{\sigma}(\mathbb{R}^d) = B_{\infty,\infty}^{\sigma}(\mathbb{R}^d) = C^{(\sigma)}(\mathbb{R}^d)$. However, the exact smoothness norm we employ is the $B_{\infty,\infty}^{\sigma}$ norm, defined in terms of wavelet coefficients. Smooth functions can be expanded as $f = \sum_{j=0}^{\infty} c_j \psi_j$ and the smoothness seminorm is expressed in terms of coefficients c_j .

A totally standard construction, used also in [3], indexes wavelets by gendered, dyadic cubes: $\nu \in \mathcal{D}$, where each $\nu = (e_{\nu}, I_{\nu})$ is a pair comprising:

- a *gender* $e = e_{\nu} \in 0, 1^d \setminus \{0\}$
- and a *dyadic cube* $I := I_{\nu} = 2^{-j}(k + [0, 1]^d)$.

For a general dyadic cube of this form we denote the corner by $c(I) := 2^{-j}k$ and the sidelength by $\ell(I) := 2^{-j}$. These definitions extend for gendered cubes: $c(\nu) := c(I_\nu)$ and $\ell(\nu) := \ell(I_\nu)$. Under this indexing, each gendered cube ν has exactly one parent ν' , where $I_\nu \subset I_{\nu'}$, $e_\nu = e_{\nu'}$ and $\ell(\nu') = 2\ell(\nu)$.

The wavelet system we employ is a family of C^r , compactly supported functions, with $r > 2k$. Each wavelet is related to one of $2^d - 1$ prototypes by affine changes of variable: $\psi_\nu(x) = \Psi_{e_\nu} \left(\frac{x - c(\nu)}{\ell(\nu)} \right)$. In other words, each wavelet is a translated, rescaled copy of one of the $2^d - 1$ functions $\Psi_e \in C^r(\mathbb{R}^d)$. Consequently the supports of wavelets are obtained by affine transformations, and each is contained in a ball with radius proportional to the sidelength and centered at the corner of the cube $I(\nu)$. I.e., there is $\Gamma > 0$ so that for all $\nu \in \mathcal{D}$

$$\overline{I}_\nu := \text{supp}(\psi_\nu) = c(\nu) + \ell(\nu) \times \text{supp}(\Psi_{e_\nu}) \subset B(c(\nu), \Gamma \ell(\nu)).$$

For orthogonal wavelet systems, compactly supported continuous functions have the unique expansion $f = \sum_{\nu \in \mathcal{D}} f_\nu \psi_\nu$. The smoothness seminorm of f is

$$|f|_{B_{\infty,\infty}^\sigma} := \sup_{\nu \in \mathcal{D}} (\ell(\nu)^{-\sigma} |f_\nu|).$$

Theorem 5 (DeVore Ron II). *Let $\ell > 2k - d + 1$. Suppose that ρ satisfies Slow Growth with $\epsilon > \frac{2k}{\ell}$. There is $C > 0$ so that for all compactly supported $f \in B_{\infty,\infty}^\sigma$, $\sigma < 2k$, there is $s_{f,\Xi} \in \text{span}(\phi, \Xi)$ so that*

$$|f(x) - s_{f,\Xi}| \leq C \rho(x)^\sigma \|f\|_{B_{\infty,\infty}^\sigma}.$$

Proof. We split $f = g + b$ where $|b(x)| \lesssim \rho(x)^\sigma |f|_{B_{\infty,\infty}^\sigma}$ and $|\Delta^k g(x)| \lesssim \rho(x)^{\sigma-2k} |f|_{B_{\infty,\infty}^\sigma}$. A consequence of this and Theorem 4 is that $|f(x) - T_\Xi g(x)| \lesssim \rho(x)^\sigma \|f\|_{B_{\infty,\infty}^\sigma}$ and the theorem follows with $s_{f,\Xi} = T_\Xi g$.

To obtain the split, partition $\mathcal{D} = \mathcal{D}_g \cup \mathcal{D}_b$ by selecting cubes ν according to the density over \overline{I}_ν :

$$\nu \in \mathcal{D}_g \quad \text{iff} \quad \ell(\nu) \geq \rho(\nu) := \max_{y \in \overline{I}_\nu} \rho(y).$$

Define $g := \sum_{\nu \in \mathcal{D}_g} f_\nu \psi_\nu = \sum_{\ell(\nu) \geq \rho(\nu)} f_\nu \psi_\nu$. Estimating the iterated Laplacian of a wavelet is straightforward: for $x \in \overline{I}_\nu$, $|\Delta^k \psi_\nu(x)| \leq C \ell(\nu)^{-2k}$. Consequently,

$$|\Delta^k g(x)| \leq C \sum_{\substack{\nu \in \mathcal{D}_g \\ x \in \overline{I}_\nu}} f_\nu \ell(\nu)^{-2k} \leq C |f|_{B_{\infty,\infty}^\sigma} \sum_{\substack{\nu \in \mathcal{D}_g \\ x \in \overline{I}_\nu}} \ell(\nu)^{\sigma-2k} \leq C' |f|_{B_{\infty,\infty}^\sigma} (\rho(x))^{\sigma-2k}.$$

The final estimate deserves some explanation. Note that $\nu \in \mathcal{D}_g$ and $x \in \overline{I}_\nu$ imply that $\ell(\nu) \geq \rho(x)$. Finding $j \in \mathbb{Z}$ so that $2^j \geq \rho(x) > 2^{j-1}$, the number of wavelets with $\ell(\nu) = 2^j$ that have x in their support is bounded, $\#\{\nu : \ell(\nu) = 2^j, x \in \overline{I}_\nu\} \leq N$, with a constant independent of x and j . We may rewrite the sum in the next to last expression in the chain of inequalities as

$$\sum_{\substack{\nu \in \mathcal{D}_g \\ x \in \overline{I}_\nu}} \ell(\nu)^{\sigma-2k} \leq \sum_{\substack{\ell(\nu)=2^j \\ x \in \overline{I}_\nu}} \sum_{i=0}^{\infty} (2^{j+i})^{\sigma-2k} \leq N 2^{j(\sigma-2k)} \sum_{i=0}^{\infty} (2^i)^{\sigma-2k} \leq C (\rho(x))^{\sigma-2k}.$$

Estimating the size of $b(x) := \sum_{\nu \in \mathcal{D}_b} f_\nu \psi_\nu$, we write $|b(x)| \leq \sum_{\nu \in \mathcal{D}_b, x \in \bar{I}_\nu} |f_\nu|$, which is bounded by $|f|_{B_{\infty,\infty}^\sigma} \sum_{\nu \in \mathcal{D}_b, x \in \bar{I}_\nu} (\ell(\nu))^\sigma$. If x and y are in \bar{I}_ν and if $\rho(y) > \ell(\nu)$ then

$$\rho(x) \geq C_{sm} \rho(y) \left(1 + \frac{|x-y|}{\rho(y)}\right)^{-r} \geq \ell(\nu) C_{sm} (1+2\Gamma)^{-r} \Rightarrow \ell(\nu) \leq C \rho(x)$$

As in the case of g , it follows that the series $|b(x)| \leq |f|_{B_{\infty,\infty}^\sigma} \sum_{\ell(\nu) \leq C \rho(x), x \in \bar{I}_\nu} (\ell(\nu))^\sigma$ can be rewritten as a sum of geometric series, to obtain $|b(x)| \leq C |f|_{B_{\infty,\infty}^\sigma} (\rho(x))^\sigma$. \square

3 Discussion

We now turn to a discussion of how condition (3) may be directly implemented to produce effective global approximants with local error estimates.

Remark 1. It is possible to generate a set of centers with a fixed spacing, having a more refined spacing on a particular subset. This can be done in such a way that the pointwise error from surface spline approximation reflects the local arrangement of centers.

In this example the global spacing is $h = 2^{-j}$, while the spacing near the origin will be $h^2 = 2^{-2j}$. We choose to impose a slow growth condition with $\epsilon = 1/3$. By Theorem 4 we observe that the LD must have precision $\ell > 2k/\epsilon$, so we choose $\ell = 7k$. The spacing of centers immediately around the origin should be 2^{-2j} , but there is an intermediate region where the spacing grows slowly. We decompose this in j regions of increasing width:

$$\Omega_J := \{x : |x| \leq 7k \times 2^{\frac{3J}{2}-2j}\} \setminus \bigcup_{0 \leq \ell < J} \Omega_\ell, \quad J = 1, \dots, j$$

in which we place gridded centers with spacing $h_J = 2^{J-1-2j}$. (The presence of the factor $7k$ is only to ensure that an adequate unisolvent set is captured in each Ω_J . Since the radii of these sets are shrinking more slowly than the spacing of the centers, this factor can be omitted for sufficiently large values of j .) For this distribution, and for points x in the transition region $\bigcup_{J=1}^j \Omega_J$ we have $\rho(x) = \rho(0)(1 + \frac{|x|}{\rho(0)})^{2/3}$. (It is not hard to show that $\rho(x) \leq \rho(y)(1 + \frac{|x|}{\rho(y)})^{2/3}$ holds throughout \mathbb{R}^d). Thus, by Theorem 5 there is a constant C so that for any f with smoothness $0 < s \leq 2k$ we have a good approximant with dramatically increased accuracy at the origin:

$$\|f - s_f\|_\infty \leq C h^s \|f\|_s \quad \text{and} \quad |f(0) - s_f(0)| \leq C h^{2s} \|f\|_s$$

In fact, for $x \in \Omega_J$, one has $|f(x) - s_f(x)| \leq C 2^{s(J-2j)} \|f\|_s$. The cardinality of centers in the ball $B(0, 7k \times 2^{-j/2})$ is less than

$$\sum_{J=0}^j (7k)^d 2^{d(\frac{3J}{2}-2j)} 2^{-d(J-2j)} \leq (7k)^d \sum_{J=0}^j 2^{d(\frac{J}{2})} \leq C_d (7k)^d 2^{d(\frac{j}{2})}.$$

Since the same ball filled with centers having uniform spacing $h = 2^{-j}$ holds roughly $(7k)^d 2^{d(j/2)}$ centers, the increased precision comes at a cost of adding only a multiple of the original centers.

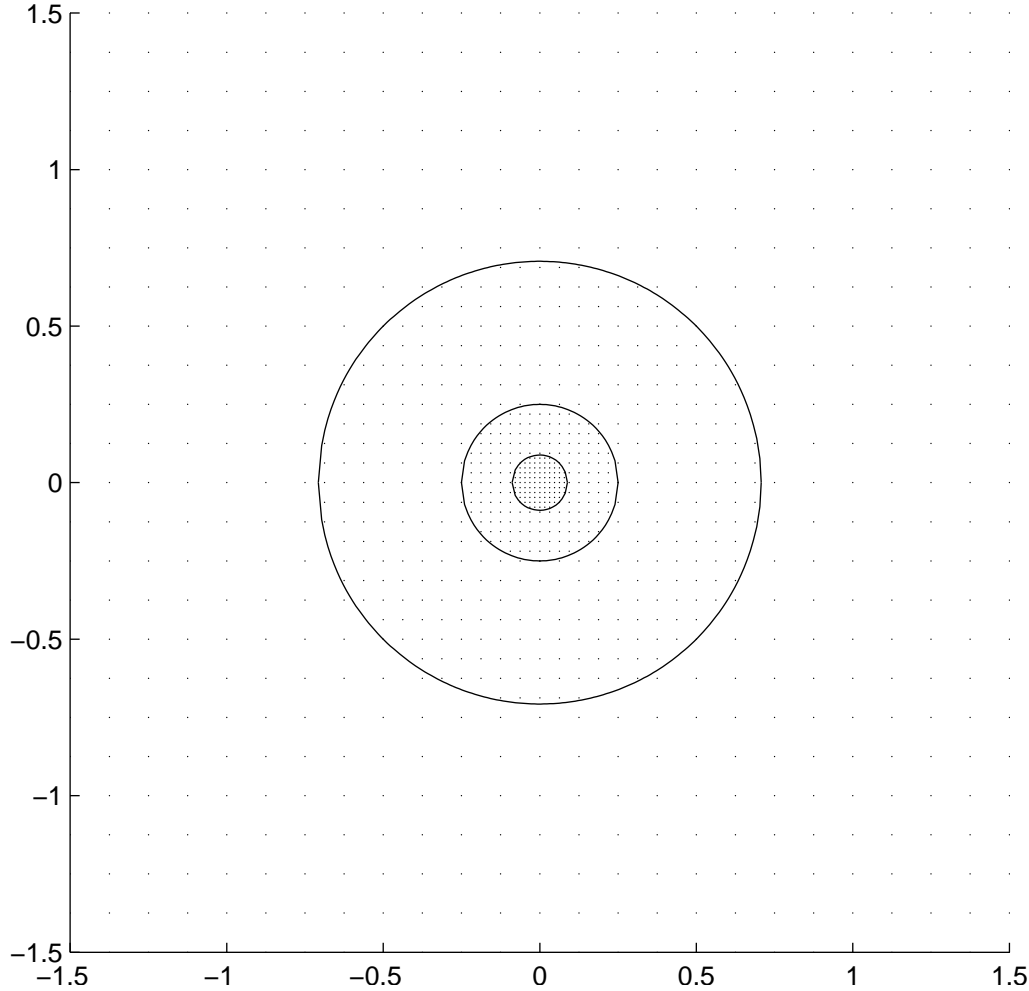


Figure 1: A configuration of centers for thin plate spline approximation ($k = 2$, as in the first remark) where the density ranges from $\rho(0) \sim h^2 = 2^{-6}$ to the coarsest density is $\rho(x) \sim h = 2^{-3}$.

We note that this example could easily be modified to place higher density centers on any subset X (curves, hypersurfaces, etc.). To do so, one produces rings Ω_j around X and fills each ring with gridded centers having density 2^{J-2j} as before.

Remark 2. In the previous remark, the spacing of centers was varied over an intermediate region having radius roughly $h^{1/2}$. The size of this region is related to the choice of ϵ in the slow growth condition, and ultimately to the degree of polynomial precision for the local polynomial reproduction. For a similar set up, where a density of roughly h^s is desired in a subset, we note that the radius of the transition region is determined by the inequality

$$h \leq h^s \left(\frac{|x|}{h^s} \right)^{1-\epsilon} \Rightarrow h^{\frac{1-s\epsilon}{1-\epsilon}} \leq |x|,$$

which suggests that $s\epsilon$ should be less than 1 in order for the transition region to stay manageable (i.e., in order that it shrinks with h). In turn, this indicates that a polynomials of a sufficiently high order should be reproduced. Thus, for given s , $\epsilon < 1/s$ and the degree of polynomial precision $\ell > 2k/\epsilon > 2ks$.

Remark 3. It should be noted that the “power function” method for estimating error from RBF interpolation, [4, Chapter 11], also provides local error estimates for RBF interpolation. The estimate on the power function [4, Theorem 11.9] can easily be rephrased in terms of a LD ρ of sufficient precision. For surface spline interpolation, the requirement is precision $k - 1$.

Remarkably, no extra global compatibility assumption such as Definition 2 are needed for this. The pointwise estimate [4, Theorem 11.16] can be stated as

$$|f(x) - s_{f,\Xi}(x)| \leq C\rho(x)^{k-d/2}|f|_{W_2^k(\mathbb{R}^d)}$$

for $f \in W_2^k(\mathbb{R}^d)$.

The drawback of this result is that the target function must reside in the space $W_2^k(\mathbb{R}^d)$, rather than permitting error estimates for the entire scale of smooth functions $C^{2k}(\mathbb{R}^d)$ or $B_{\infty,\infty}^\sigma(\mathbb{R}^d)$, $\sigma < 2k$ as in Theorem 4 and Theorem 5.

References

- [1] C. DE BOOR AND G. J. FIX, *Spline approximation by quasiinterpolants*, J. Approximation Theory, 8 (1973), pp. 19–45. Collection of articles dedicated to Isaac Jacob Schoenberg on his 70th birthday, I.
- [2] R. A. DEVORE AND G. G. LORENTZ, *Constructive approximation*, vol. 303 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1993.
- [3] R. A. DEVORE AND A. RON, *Approximation using scattered shifts of a multivariate function*. Unpublished manuscript.
- [4] H. WENDLAND, *Scattered data approximation*, vol. 17 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2005.